International Journal of Environment, Agriculture and Biotechnology
Vol-10, Issue-2; Mar-Apr, 2025
Peer-Reviewed International Journal
Journal Home Page Available: https://ijeab.com/
Journal DOI: 10.22161/ijeab

Genome-Wide Identification and Characterization of Superoxide Dismutase (*SOD*) Gene family in Finger Millet (*Eleusine coracana*)

Viswanadha Naik Jarapala[1,2], Anjana priyadarshani Kanathala[1,2], Vikas Reddy Olathru[2], Prashanth Bollempally[2], Vijay Kumar G[2], Srinivas Naik Kethavath[1*]

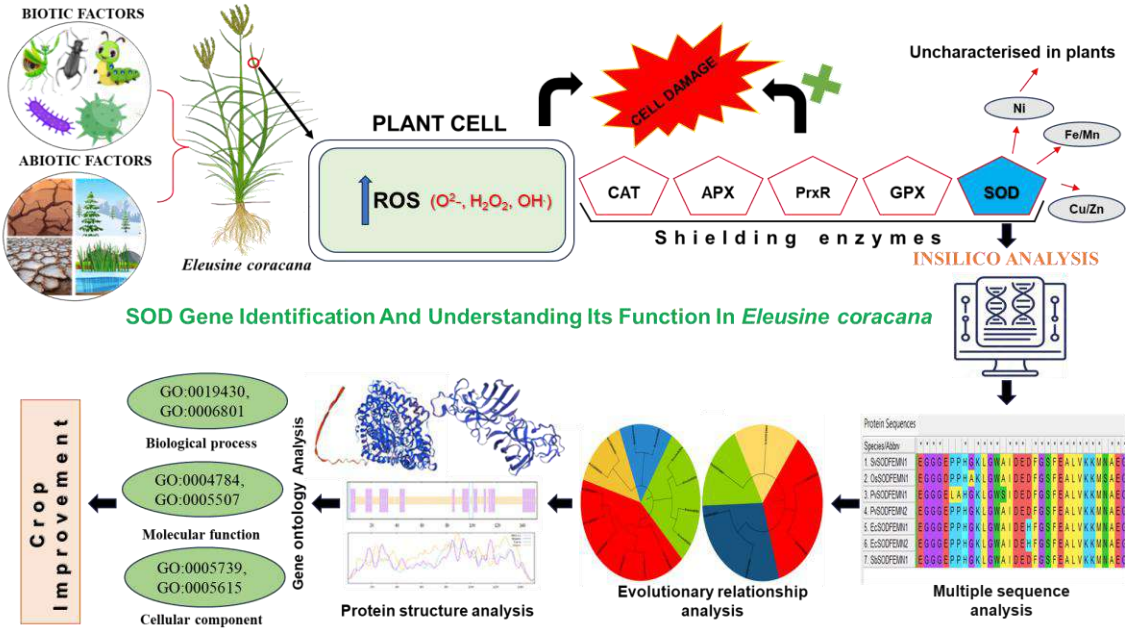
[1]Centre for Plant Molecular Biology, Osmania University, Hyderabad -500 007, India
[2]Department of Genetics and Biotechnology, Osmania University, Hyderabad -500 007, India
*Corresponding author: srinivasnaik.cpmb@gmail.com

Received: 14 Oct 2024; Received in revised form: 13 Dec 2024; Accepted: 19 Mar 2025; Available online: 26 Mar 2025

Abstract— *The metallo-enzyme superoxide dismutases (SODs) are significant in protecting plants from environmental challenges as well as regulating their growth and development. Although many plants have been determined to possess SOD gene families in their genomes, it is known, very scarcely, about such gene families in finger millet (Eleusine coracana). This study explored the SOD gene family across the entire genome of finger millet. A total of 10 SOD genes were discovered, comprising eight Cu/ZnSODs and two Fe/MnSOD. These EcSODs are spread irregularly throughout 3A and 8B chromosomes. Phylogenetic analysis revealed that SOD proteins in plants possibly classified into three primary groups for both EcCu/ZnSODs and EcFe-MnSODs. The motif and exon/intron makeup of SOD genes are conserved within the same subgroup. Protein structure prediction showed all homologs contains highest similarity with SOD peptide structure. Furthermore, numerous cis-elements that react to distinct stresses were distributed differently. The various biological processes associated with background molecular roles of SODs are further demonstrated by gene ontology analysis. The transcriptional factors discovered indicate that SODs are mostly connected to external environmental and biotic stress. This study lays the groundwork for future cloning, Genetic manipulation of SOD gene in finger millet which contributes towards finger millet breeding programs.*

Keywords— ***Finger millet (Eleusine coracana), Superoxide dismutase (SOD), ROS, Abiotic and Biotic stress.***

GRAPHICAL ABSTRACT



I. INTRODUCTION

Millets were among the initial small-seeded nutri-cereals annual grass species to undergo domestication. Their exceptional nutritional value with climate-resilience properties enables them to serve as a traditional food source in many countries like Africa and Asia. These staple grains are a part of daily routine in most of the developed and under developed countries and consumed by more than 590 million people. As more people seeking healthy diet options, the global market of millet is projected to grow, potentially reaching $12 billion by 2025 (1). India contributes 11.42 (37%) million tons of the 30.73 million tons of millets produced worldwide currently and sustaining its place as the top producer (2). Originating in the highlands of Ethiopia, finger millet is now extensively cultivated in over 25 nations (3). *Eleusine coracana L. Gaertn.* is regarded as an orphan cereal (4). which is used for food as well as fodder. It is also known as hardy crop due to its widespread farming in barren and harsh environments as well as in poor, dry, and less fertile soils. Thus, finger millet serves as a boon in areas of extreme poverty (5,6,7). Finger millet is said to be more nutritious than wheat, rice, and maize because its grains are high in dietary fiber, proteins, minerals, vital amino acids, vitamin B complex, iron, and calcium and many more, (5). It is free from gluten and provides an added benefit for those with digestive problems. It also offers a wide range of health benefits, which includes anti-diabetic (type 2 diabetes mellitus), anti-diarrheal, antibiotics, anti-allergic, atherosclerogenic, anti-ulcer, antitumorigenic (for K562 chronic myeloid leukemia), and antioxidant qualities (8).

Finger millet is constantly subjected to environmental stresses which hampers its yield. Abiotic stress includes heat, cold, drought, salt, metal stress, ozone, UV radiation, and nutrient deficiencies (9,10). Biotic stress includes following diseases- blast, foot rot, and pests like pink stem borer and root aphid (11). An especially destructive disease that significantly reduces finger millet output is *Magnaporthe grisea* (anamorphic stage: *Pyricularia grisea*), an ascomycete filamentous fungus that causes finger millet blasts (12). Blast causes yield losses ranging from 7.32% to 90% in finger millet depends on which tissue it occurs (13). A yield loss of 28% to 36% has been observed in India (14), and up to 80% elsewhere (15). To ensure global food security, there is high need to improve the genetic resources of finger millet. The most efficient, environmentally friendly, sustainable, and farmer-beneficial approaches to handle this finger millet yield loss is by using genetic manipulation and genomic improvement techniques based on deep study on interactions between plant and stress.

Whenever a plant encounters a stress, excessive accumulation of Reactive oxygen species (ROS), byproducts of cell metabolism and specialized with oxidizing properties which results in release of hydroxyl radical (OH), and superoxide (O^{2-}), along with hydrogen peroxide (H_2O_2). When ROS levels are appropriate, organisms require it. It can influence several plant physiological processes in various species by acting as a signalling molecule (16). Overproduction of ROS can cause membrane lesions, metabolic disruption, and even cell death. It can also harm biological macromolecules lipids,

proteins, and nucleic acids (17). Plants have evolved effective defense mechanisms against ROS induced damage in order to combat its toxicity by means of elimination of surplus of ROS. Some shielding enzymes, including superoxide dismutase (SOD), peroxiredoxin (PrxR), catalase (CAT), glutathione peroxidase (GPX), and ascorbate peroxidase (APX), can keep the dynamic equilibrium of ROS levels in plants (18). Among the mentioned enzymes, *SODs* are mentioned as the primary enzyme involved in plant defense system and have the potential to mitigate ROS-induced damage by catalysing the breakdown of superoxide radicals into H_2O_2 and O_2 (19). They mitigate the risks associated with ROS by forming molecular oxygen (O_2) and hydrogen peroxide (H_2O_2) from superoxide (O_2^-) in a oxidative stress (20). As metalloenzymes, *SODs'* peptides need metal cofactors in order to function as catalysts. Plant *SODs* are divided into three classes based on their metal cofactor which are iron *SODs (FeSODs)*, manganese *SODs (MnSODs)*, and copper/zinc *SODs (Cu/ZnSODs)* (21). Nickel superoxide dismutase (*NiSOD*), another class of *SOD*, has not been found in any plant (22).

Verma D et al, 2019 stated that *SOD* gene serves its critical functions in plants by responding to biotic and abiotic stress. Certain research has looked at sea-grasses unique reaction to oxidative stress and the responses of *SOD* in hot conditions (24). It has been demonstrated that the *ZmMnSOD* enzyme is crucial in reducing oxidative damage response during temperature stress (25). In recent study by Madanala et al. (2011) (26) it was found that *Withania somnifera* plant species contains a highly stable *Cu/Zn-SOD* gene in their chloroplast and this gene shows a greater resistance to ethanol and detergent (27) like external substances. *MnSOD's* higher expression in tomatoes confers resilience against salt and oxidative damage (28). Advances in sequencing technologies uncovered multiple activities of *SOD* genes and led to the identification of *SOD* gene families across the complete genomes of different plant species (29).

Finger millet has a high degree of genetic diversity in terms of agronomic features, dietary value, and root properties (30). Hence, genomic techniques and databases are critical for phenotypic exploration, gene mining, and marker-assisted breeding (31). Transcriptomics has been utilized in finger millet to annotate the genome, identify markers, and locate candidate genes (32). Genome assets are most important tools for designing high yielding stable finger millet cultivars and increasing genetic diversity in response to changing environmental circumstances.

The present analysis is a comprehensive genome-wide analysis of the *SOD* gene family in *Eleusine coracana* and compared the findings with previous findings of *Arabidopsis thaliana* in order to explore the significance, functionality, and evolutionary relationships of the *SOD* gene family.

II. MATERIALS AND METHODS

IN-SILICO EcSOD GENE IDENTIFICATION, AND CHARACTERIZATION

EcSOD genes were retrieved from finger millet using protein sequences obtained from phytozome database (http://www.phytozome.net/) (33). *Arabidopsis's AtSOD* gene was utilized to BLAST P against the *Eleusine coracana* protein database. The hits were selected based on a threshold E value of <1E-5. The selected genes were provided to the motif finder tool, which revealed the domain in protein sequences. The genes containing the SOD Domain were selected from the findings, whereas others were declared redundant.

GENE STRUCTURE PREDICTION, MOTIF STUDY and SUB CELLULAR LOCALIZATION

Eleusine coracana GTF data and Gene Structure Display Server 2.0 were employed to determine structural patterns of exons and introns. (https://gsds.gao-lab.org/index.php) (34). To predict conserved motifs, a set of parameters were applied: a maximum of 10 motifs, 2–20 motif positions, and 6–20 widths using MEME (http://tools.meme-suite.org/) (35). The conserved domain analysis was done using NCBI-CDD search. WOLFPSORT is used for protein Subcellular Localization Prediction. (https://www.genscript.com/wolf-psort.html) (36).

PHYLOGENETIC, SYNTENY AND PROMOTER/ CIS-ELEMENTS ANALYSIS

The amino acid sequences of *SOD* genes from Rice and Sorghum were aligned using Mega version 11.0. (megasoftware.net). Mega-program was used to create the phylogenetic genetic tree using a neighbour joining technique and a thousand bootstrap replications (5). Paralog genes synonymous and non-synonymous ratios were computed using the Ka/Ks calculator. (KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. Genomics Proteomics Bioinformatics 2021). The candidate *EcSOD* gene homologs were mapped to *S. bicolor* and *O. sativa* genomes, and synteny maps were developed using TB tools. The multiple circos similarity analysis was performed using circoletto online server. (http://bat.infspire.org/tools/circoletto) through bit score and % identity methods.

The phytozome database included the promoter sequences for every gene that was chosen for further analysis. These sequences are 1500 bases upstream from the matching genomic sequence of each gene. With the use of these downloaded promoter sequences, the PLANT CARE tool anticipates CIS regulatory components. (http://webtools/plantcare/html/bioinformatics.psb.ugent.be).

PHYSICAL MAPPING OF SOD GENES ON CHROMOSOMES

The physical mapping of genes on finger millet chromosomes was done through Phenogram tool. (https://visualization.ritchielab.org/phenograms/plot) The results were downloaded with exception of empty chromosomes.

GENE ONTOLOGY, TRANSCRIPTION FACTOR ANALYSIS and CpG Island PREDICTION -

Since PTFDB plant transcription factor database(http://gao-lab.org/planttfdb/) does not have data on finger millet, it was necessary to pick closely similar species, such as *Seteria viridis*, in order to estimate the connection sites of all *EcSOD* protein transcription factors. With the aid of these data, network was created with the Cytoscape application (https://cytoscape.org). Using Cytoscape Tool, the data are captured and shown as a network of *EcSOD* genes. The homologous gene ontology has been analyzed using PANNZER2.0. The upstream 200-2000bps promoter sequence was taken with default parameters. CpG islands for EcSOD homologs were identified using methprimer tool (MethPrimer-Design MSP/BSP primers and predict CpG islands - Li Lab, PUMCH (54). The promoter regions 200-2000 bp upstream sequences as well as gene body sequence was taken for analysis.

PROTEIN STRUCTURE PREDICTION, PROTEIN PARAMETERS and PPI INTERACTIONS

Using the Swiss model server a fully automated protein structure homology modelling server, the 3D protein structures of the *EcSOD* transcripts were estimated. (https://expasy.org/swissmodel/) Using Ramachandran plots, the PSVS protein structure verification site (https://saves.mbi.ucla.edu/) is used to verify and analyze the stability of predicted protein structures. The 2D structure of *EcSOD* homologs was identified using SOPMA server (37).

(https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html).

The STRING database was used to predict and observe the PPI protein-protein interactions of these homologs. The URL is https://string-db.org. The protParam program predicts properties of proteins, including the isoelectric point, gravy (grand average of hydropathy), and aliphatic indices.

INSILICO EXPRESSION ANALYSIS-

The expression data of *EcSOD* genes under drought stress were retrieved from Milletdb (http://milletdb.novogene.com/home/) database and TKM values of control and treated plants were mapped by TB tools.

III. RESULTS

GENE - IDENTIFICATION AND CHARACTERIZATION.

The sequence from *Arabidopsis thaliana* is taken and used as query for BLAST P search against *Eleusine coracana* genome to identify of *EcSOD* genes. From all the hits obtained, we carefully chosen 10 genes out of which 8 genes contain CU-ZN domain and 2 genes contains FE-MN domains. These genes were subjected to motif Finder tool for identification of SOD domain. We got 10 sequences containing SOD domain which are renamed as *EcSODCUZN1, EcSODCUZN2, EcSODCUZN3, ……. EcSODCUZN8,* and *EcSODFEMN1, EcSODFEMN2* with regard to their ligand and location on pseudo chromosomes. (Table1). Gene characterization was done and proteins with longest peptide sequence were *EcSODFEMN1* and *EcSODFEMN2* and protein with shortest were *EcSODCUZN1, EcSODCUZN5* and *EcSODCUZN6.* Analysis was done to obtain molecular weight of proteins and the outcome was all 10 *EcSOD* proteins are in between 15kDa to 21KDa. We analysed protein parameters which show that Pi ranged from 5.34 to 7.92, GRAVY ranged from -0.075 to -0.272. (Table2). Gene structure predictions revealed that among 8 *EcSODCUZN* genes, *EcSODCUZN4* and *EcSODCUZN5* has 7/8 intron to exon ratio; *EcSODCUZN1, EcSODCUZN2, EcSODCUZN3, EcSODCUZN6, EcSODCUZN7, EcSODCUZN8* has 6/7 intron to exon ratio; All the genes have no UTR regions. *EcSODFEMN1* and *EcSODFEMN2* has 5/6 intron to exon ratio, was nearly similar among same group members. (Figure 1b). The subcellular localization of *EcSODCUZN* homologs were found to be in cytoplasm for *EcSODCUZN1, EcSODCUZN2, EcSODCUZN5, EcSODCUZN6, EcSODCUZN7, EcSODCUZN8;* whereas *EcSODCUZN3, EcSODCUZN4* were in chloroplast. The subcellular localisation of *EcSODFEMN1* and *EcSODFEMN2* were in Mitochondria.

PREDICTION OF MOTIFS AND CONSERVED DOMAIN.

MEME tool was used to spot different types of motifs. Out of 10 motifs found, most of them are present among genes

and in same patterns within the same group. (Figure 8a). Results from conserved domain and motif patterns of candidate gene revealed that Super oxide dismutase domain is conserved among all genes.

The three domains were observed but domain with ID-PLN02386 also belongs to superoxide dismutase -Cu/Zn superfamily (cl00891), the protein homologs with *EcSODFEMN* showed domain specific for metal Mn/Fe superoxide dismutase. (Figure1c). The pattern of copper or zinc SOD domain was similar in all proteins. The *EcSODCUZN* genes were having extra motif of *Cu/Zn SOD*. The motif sequence pattern of these is shown in figure 8b and 8c. This identical patterns in motifs of genes among other groups relates with evolutionary relationships

MULTIPLE SEQUENCE ALIGNMENT, PHYLOGENY and SYNTENY and SIMILARITY AMONG SPECIES AND Ka/Ks ANNOTATION.

Multiple sequence alignment of *EcSODFEMN* gene homologs (Fig 2a) with protein sequences of *OsSODFEMN1, SvSODFEMN1, PvSODFEMN1, PvSODFEMN2, SbSODFEMN1 and EcSODCUZN* homologs (Fig 2b) with peptide sequences of *OsSODCUZN1, SvSODCUZN1, SvSODCUZN2, PvSODCUZN1, PvSODCUZN2, SbSODCUZN1* were analysed using Mega 11.0 by ClustalW and Visualised through Snapgene viewer, the alignment showed that there are identical frequencies, which means they are conserved among the species.

The evolutionary relationship analysis among *EcSOD* homologs exposed 5 paralog pairs with 3 groups. *EcSODCUZN1, EcSODCUZN3; EcSODCUZN5, EcSODCUZN8; EcSODCUZN3, EcSODCUZN4; EcSODCUZN2, EcSODCUZN7;* and *EcSODFEMN1, EcSODFEMN2*; are the paralog pairs formed (Fig 1a).

The phylogenetic analysis of candidate genes with closely related species like *Seteria viridis, Panicum virgatum, Sorghum bicolor, Oryza sativa* gave 4 groups with both *CUZN* and *FEMN* homologs respectively. *SvSODCUZN1, PvSODCUZN1, SbSODCUZN1, OsSODCUZN1, EcSODCUZN1, EcSODCUZN6* formed orthologous pairs with in one group. *EcSODCUZN2, EcSODCUZN3, EcSODCUZN4,* and *EcSODCUZN7* shown no orthologous pairs with other species. *EcSODCUZN5 and EcSODCUZN8; SvSODCUZN2, PvSODCUZN2*; formed into two other groups. The SODs with metal factor Fe/Mn from finger millet goes into one group forming no paralog pairs with other species (Figure 3a and 3b).

The similarity analysis by multiple circos among species was performed and the results were downloaded. Red colour (Fig 9a) indicating higher bit score value (>0.75) between the genes compared to others and orange colour (Fig 9b) indicated higher %identity (99.99%) among protein sequences. The *EcSOD* genes shown similarity with *Sorghum bicolor* and *Seteria viridis* species. Three paralog pairs are identified from Ka/Ks annotated evolutionary tree analysis results. All paralogs of *EcSOD* showed Ka/Ks c ratio not more than 0.21 (<1) which signifies specific purifying selection. The tabulated Ka/KS values and paralog pair given in Table3. Ka/Ks ratios of >1, =1, and <1 indicate positive, purifying, and neutral evolution, respectively.

Synteny analysis with *S. bicolor* shown candidate genes were distributed across the genome (Fig 2c). The genes on chromosome 8A and 8B of *E. corcana* shows homology with chromosome 07 of *S. bicolor.* chromosome 5A, 5B of *E. corcana* shows homology with chromosome 09 of *S. bicolor.* the genes distributed on *E. corcana* chromosome 3A, 3B collinear with chromosome 01 in *S. bicolor*. Genes on chromosome 7A and chromosome 7B of *E. corcana* shows collinearity with S. bicolor chromosome 01. Results of synteny between *E. corcana and O. sativa* revealed chromosome 8A, 8B similarity with chromosome 08; chromosome 5A, 5B with chromosome 05; chromosome 3A, 3B with chromosome 03; chromosome 3A, 3B, 7A, 7B with chromosome 07 respectively (Fig 2d). Collinearity was observed in between *M. sinesis* and *E. corcana.* Chromosome 8A, 8B homologs with chromosome 03, 07 on *M. sinesis*. Chromosome 5A, 5B homologs with chromosome 16, 17 on *M. sinesis.* Chromosome 3A, 3B homologs with chromosome 01 on *M. sinesis.* Chromosome 02 of *M. sinesis* has similarity for genes on chromosomes 3A, 3B, 7A, 7B of *E. corcana* (Fig 2e). There is no collinearity among *E. corcana* and *S. viridis*.

Physical Mapping, Cis-Elements Prediction, and their Distribution -

Transcription is a crucial step to start the production of genes and it is also a point where RNA polymerase binds with promoter like regulatory regions. The structure of the promoter is essential for RNA polymerase binding affinity, which in turn affects the degree of gene expression (38). Promoter analysis of *EcSOD* homologs revealed the occurrence of Abiotic stress, light responsive and Phyto-hormone responsive putative cis-regulatory elements (Fig 7b). The frequency of occurrence of cis elements were tabulated and presented as a heatmap (Fig 7a). Light responsive elements and MEJA were highly distributed among the genes. MYB- Related, ABA responsive elements are moderately present in most of them. The regulatory regions of *EcSOD* gene homologs contain a variety of cis-regulatory elements, including those associated to MYB-flavonoid genes, MYB-drought responsive, MYB-light responsive, and defense responsive. MEJA, GA, SA, and

AR are phytohormone-responsive elements found in the majority of *SOD* homologs. Across the Finger millet genome, the locations of *EcSOD* homologs were mapped (Fig 10a).

PROTEIN MODELING- 2D,3D AND PROTEIN-PROTEIN INTERACTIONS (PPIs)

The Swiss model is used here which predict 3D structure of various proteins based on PDB structures. High similarity models were selected from standard structures available in database. The appropriate template ID, protein ID, percentage identity were noted and tabulated (Table 2). 3D structures are given in (Fig 4).

The results of secondary structure were tabulated in Table 4. Beta turn percentage was zero in all the *SOD* proteins. In case of *EcSODFEMN* homologs the % of alpha helices was higher approx. 50% compared to *EcSODCUZN* homologs. The structure of *EcSODCUZN*1 and *EcSODFEMN1* are saved (Fig 5a, 5b).

The structure ID B4F925.1.A with superoxide dismutase from mitochondria, showed structure similarity with *EcSODFEMN* homologous. Structure ID 3Km2.3. A Superoxide dismutase with copper/zinc from chloroplast region, Shown similarity with *EcSODCUZN1, EcSODCUZN5, EcSODCUZN6, EcSODCUZN8*. Structure P93407.1.A – SOD (Cu-Zn), Showed similarity with *EcSODCUZN3, EcSODCUZN4. EcSODCUZN2* shown highest similarity with K4AFE1.1. A – SOD (Cu/Zn). The structure similarity ranged between 64 to 94% among homologs.

A network displaying both direct and indirect links to the candidate proteins was built using an online database-String database. The interpretation of the results reveals interactions with A0A368QPZ2, A0A368RQ49, A0A368SFS1, A0A368STV7, K3YIU7_SETIT, K3Z9A6_SETIT, K3ZXQ6_SETIT, and K4AFE1_SETIT, which are primarily involved in the metabolism of ROS, carbohydrates, the glyoxal metabolic pathway, and superoxide dismutase activity within cells in response to oxidative stress (Figure 5c).

TRANCRIPTIONAL FACTORS PREDICTION-

The results of transcription factor analysis showed that 35 different kinds of transcriptional factors are associated with candidate genes. Abiotic stress and biotic-related transcriptional factors, such as MYB, WOX, EIL, TALE, Trihelix, CAMTA, C2H2, LBD, ERF, MYB-related, HD-zip, BZIP, GATA, SBP, MIKC-MADS, AP2, NAC, WRKY, TCP, bHLH, and G2-like, are primarily engaged in plant growth and development. (Figure 6). (39), (40). Certain transcriptional factors, such as Dof and GATA, are associated with growth and development of plants, while some are hormone-responsive Tfs.

INSILICO EXPRESSION ANALYSIS –

Expression data reveals that *SOD* gene levels increased under drought stress compared to untreated ones in both the cultivars (IE7079, IE6537). Among the homologs *EcCUZNSOD1, EcCUZNSOD3, EcFEMNSOD1* were highly expressed under drought treatments. Results show that *SOD* is upregulated under stress and a stress responsive gene (Fig 10b).

OUTCOME OF CPG ISLANDS PREDICTION AND GO ANALYSIS -

Methylated regions in promoter and gene body were identified and results were tabulated (Table 6). *EcCUZNSOD2, EcCUZNSOD5, EcCUZNSOD8, EcFEMNSOD1* showed three CpG Islands in their promoter regions. Maximum number of CpG rich sites were found in *EcCUZNSOD7* whereas *EcCUZNSOD1, EcCUZNSOD4, EcCUZNSOD6* shown no CpG rich regions in the promoter regions. *EcCUZNSOD3, EcFEMNSOD2* got two CpG islands. *EcSODCUZN* homologs are mainly involved in superoxide metabolic activity, response to ozone, oxidant detoxification and major responses in according to abiotic, salt, high intensity light, metal responses whereas *EcSODFEMN* homologs are involved in removal of superoxide radicals, oxidative stress, response towards xenobiotics and herbicides. GO in molecular function shows copper ion binding for *EcSODCUZN* homologs but for *EcSODFEMN* its metal ion binding. The cellular components for *EcSODCUZN* homologs were cytoplasm, chloroplast, and peroxisomes whereas *EcSODFEMN* located in mitochondria. The results of GO analysis come in accordance with results of prediction of subcellular localisation for the *EcSOD* protein homologs.

IV. DISCUSSION

Environmental stress causes a significant challenge to plants, leading to troubles in both morphological and physiological growth processes, toxic ROS are frequently produced in plants upon interaction with these stresses. Overexposure to ROS can cause membrane lipids, DNA strand breaks, and enzyme inactivation (41). Members of the SOD family play a preliminary plus vital role in protecting against ROS. All creatures that exist in the presence of O2 are assumed to include the enzyme antioxidant *SODs*, which disproportionately convert reactive O2•− in to H2O2 and O2 upon interactions with co-factors like Cu, Fe, or Mn. (42). Because the active core of the enzyme reaction cycle involves the interaction of two O2•− free radicals and the short-term retention of them, two

redox-active metal ions are required as cofactors (43). Most of the previous studies of plant species follow a traditional methodology such as Blast search and pfam search of known proteins of related families to identify SOD gene family target species (23). It is crucial to characterize the SOD gene family in finger millet because to its notable resistance to Environmental stresses.

Ten *SOD* genes total—eight *Cu/Zn-SODs* and two *Fe/Mn-SODs*—covering the two main categories of *SOD* genes were found in finger millet in the current study (Table 1*).SOD* gene number varies from plant to plant and various previous studies mentioned that Arabidopsis have 8 *SODS* (3 *Cu/Zn-SODs*, 2 *Mn-SODs*, and 3 *Fe-SODs)* whereas sorghum contains 8 (5 Cu/Zn-SODs, 1Mn-SOD, and 2 Fe-SODs) and tomato having 9 *SODS (4 Cu/Zn-SODs, 1 Mn-SOD, and 4 Fe-SODs*).This difference in number of SOD genes in each plant may be due to the differences in their genome size, evolutionary divergence, and Environmental Adaptation. However, not only limited to these. Gene duplication, which includes tandem and sectoral duplications are crucial for the growth of *SOD* gene diversity, might be the root cause of these variations (44), (45).

Phylogenetic investigation showed a close association among *Cu/Zn-SODs* and *Fe-SODs/Mn-SODs* members. Based on the bootstrap values, relative phylogenetic analysis of *SOD* proteins of finger millet and other crops or plants (Sorghum, Panicum, Seteria, and rice) mutually formed four distinct groups in both cases; these findings are compatible with earlier research; Regarding the subcellular location of *SODs*, most of the data supported the evolutionary conclusions. It was anticipated that the cytoplasm, mitochondria, and chloroplast would contain *SODs.* According to prior research (Fink and Scandalios 2002), the majority of cytosolic and chloroplast genes include seven introns, and intron–exon arrangements of plant *SO*D genes were shown to be well conserved (21). Furthermore, similar sequences for the majority of *EcSODs* were discovered through phylogenetic analysis of other *SOD* species, indicating that *EcSODs* most likely serve the same purposes as *SODs* in other plant species. According to earlier research, *SOD* protein clusters may be related to the subcellular placements of *SODs;* in our investigation, individuals that grouped within the same subgroup likely to have similar subcellular localizations (46, 47).

Seven intron regions have been found among the ten *EcSOD* genes through gene structure analysis. In a previous study it was revealed that *SOD* genes in plants had highly conserved introns while chloroplast *SOD* as well as most of the cytosolic *SODs* retained seven introns (21).

The three main mechanisms that may be causing this discrepancy are gain/loss of exon/intron, insertion/deletion, and exonization/pseudo-exonization according to recent research. Their enzymatic activity and expression pattern that adapts to different stress circumstances may be impacted by structural divergences (48). The examination of conserved motifs in *SOD* proteins corroborates the evolutionary information (Fig. 2). Similar concepts were shared by the same grouping. Interestingly, *SOD* proteins that were concentrated in each subgroup found to have similar motif distributions, dimensions, and locations (48).

The protein structure predicted aligns with phylogenetical relationship of candidate proteins, paralog proteins are having similar 3D structure with identical percentages of alpha helices, beta turns and extended strands of them (Fig 4 and Fig5a,5b).

A complex regulatory mechanism is required to control gene expression in response to different abiotic and biotic stress. Understanding the transcription factors and cis-elements present in the promoter sequences provides information on how *SiSODs* are regulated upstream. Our findings demonstrate that a large number of cis-elements linked to stress-responsive events were found in the promoters of *SiSODs*. MYB is a class of transcription elements that has been identified to be involved in control of physiological metabolism, organogenesis, cell morphogenesis, and growth and development in plants (49). Furthermore, biotic, and abiotic stress responses in plants were linked to several *MYB* genes (50). A large portion of *MYB TFs* helps in the establishment of host resistance against various pathogenic fungi (51). In order to protect themselves from biotic and abiotic stress, plants develop a variety of secondary metabolites. The combined activity of bHLH and MYB TFs controls the synthesis of tissue-specific flavonoids, such as phenylpropanoid (52). Plants which grow in stress free environment showed very little induction of *CAMTA* genes. This may be because the TF genes in this family have redundant functions or that the genes in this class are expressed in particular environments (53).

The cis-regulatory elements existing in the promoter regions were the binding sites of *SOD*s gene with other proteins to play an essential role in regulating gene transcription. There were a huge number of light responses associated regulatory elements, Phyto-hormone responsive elements which involves in plant defense mechanism, growth, and drought stress reactive elements. Among all the *EcSOD* homologs, MYB had the greatest number of elements, suggesting that it is associated with the production of lignin and plays a role in stress tolerance. Defense-responsive elements found in *EcSOD* homologs

suggest that these proteins are involved in defense related to biotic stress. MeJA, GA, SA, and AR are phytohormone-responsive elements found in the majority of *SOD* homologs. The defense response elements MeJA and SA have been discovered to be the most abundant among phytohormone responsive elements. They have also been confirmed to be present in all *EcSOD* homologs, suggesting that they are involved in the defense mechanism against biotic stress. The promoter regions of *EcSOD* genes also contain light-responsive elements. This discovery suggests that light might affect the Finger millet's *SOD* genes. Studies involving cis-elements are crucial as they have the potential to reveal the functional control of members of the *EcSOD g*ene family. Gene expression patterns and gene functions that are comparable between homolog sequences may be greatly influenced by similar Cis-regulatory regions. The abscisic acid-related motif ABA and the methyl jasmonate-related MEJA motif were present in a significant proportion of *EcSODs*. Different subgroups' unique regulatory elements may cause the genes in those subgroups to act differently. Expression analysis shows *SOD* gene as a stress responsive and upregulated under stress conditions in finger millet cultivars. The gene ontology studies shows that *EcSODs* are mainly related in responses to oxidative radicals, oxidative stress management and other abiotic stresses which makes *EcSOD* gene as a potential target for producing crop varieties resilient to environmental conditions and biotic infections.

Table 1: Gene characterisation of EcSOD homologs.

TRANSCRIPT ID	GENE	CHR	LOCATION START END	STRAND	Localization
ELECO.r07.5BG0424640.1	>EcFEMNSOD1	5B	11423066..11427964	Reverse	Mitochondria
ELECO.r07.5AG0377540.1	>EcFEMNSOD2	5A	13454562..13459494	Reverse	Mitochondria
ELECO.r07.3BG0279950.1	>EcCUZNSOD1	3B	51383645..51385320	Reverse	Cytoplasm
ELECO.r07.3BG0288910.1	>EcCUZNSOD2	3B	58305000..58307819	Forward	Cytoplasm
ELECO.r07.8BG0669600.1	>EcCUZNSOD3	8B	62759715..62761705	Forward	Chloroplast
ELECO.r07.8AG0641100.1	>EcCUZNSOD4	8A	47311060..47313528	Forward	Chloroplast
ELECO.r07.7AG0580610.1	>EcCUZNSOD5	7A	46693478..46694625	Reverse	Cytoplasm
ELECO.r07.3AG0248280.1	>EcCUZNSOD6	3A	48098898..48101342	Forward	Cytoplasm
ELECO.r07.3AG0239070.1	>EcCUZNSOD7	3A	40631296..40634115	Reverse	Cytoplasm
ELECO.r07.7BG0611780.1	>EcCUZNSOD8	7B	57514079..57515197	Reverse	Cytoplasm

Table 2: Protein parameters of SOD genes.

TRANSCRIPT ID	**GENE**	**A.AS**	**WEIGHT**	**PI**	**Gravy**
ELECO.r07.5BG0424640.1	>EcFEMNSOD1	236	25.51 kb	7.9	0.153
ELECO.r07.5AG0377540.1	>EcFEMNSOD2	236	25.49 kb	7.92	0.179
ELECO.r07.3BG0279950.1	>EcCUZNSOD1	152	15.13 kb	5.65	-0.139
ELECO.r07.3BG0288910.1	>EcCUZNSOD2	163	16.48 kb	6.58	-0.075
ELECO.r07.8BG0669600.1	>EcCUZNSOD3	204	20.6 kb	5.49	0.111
ELECO.r07.8AG0641100.1	>EcCUZNSOD4	204	20.6 kb	5.34	0.104
ELECO.r07.7AG0580610.1	>EcCUZNSOD5	152	15.2 kb	5.76	0.303
ELECO.r07.3AG0248280.1	>EcCUZNSOD6	152	15.13 kb	5.65	0.139
ELECO.r07.3AG0239070.1	>EcCUZNSOD7	163	16.48 kb	6.58	0.075
ELECO.r07.7BG0611780.1	>EcCUZNSOD8	152	15.20 kb	5.93	0.272

Table 3: Ka/Ks values.

Gene 1	Gene 2	Ka	Ks	Ka_Ks
>EcCUZNSOD2	>EcCUZNSOD7	0	0.074958	0
>EcCUZNSOD3	>EcCUZNSOD4	0.013499	0.0638	0.211577
>EcCUZNSOD1	>EcCUZNSOD6	0	0.072895	0

Table 4: 2D structure analysis of EcSOD proteins.

Gene	**Alpha helix %**	**Extended stand %**	**Beta turn %**	**Random coil %**
>*EcFEMNSOD1*	50.85	12.29	0	36.86
>*EcFEMNSOD2*	48.73	12.29	0	38.98
>*EcCUZNSOD1*	3.29	31.58	0	65.13
>*EcCUZNSOD2*	3.68	28.83	0	67.48
>*EcCUZNSOD3*	7.35	26.96	0	65.69
>*EcCUZNSOD4*	5.88	25	0	69.1
>*EcCUZNSOD5*	4.61	30.26	0	65.13
>*EcCUZNSOD6*	3.29	30.26	0	66.45
>*EcCUZNSOD7*	3.07	30.06	0	66.87
>*EcCUZNSOD8*	3.95	30.26	0	65.79

Table 5: Gene ontology analysis results.

GENE	**BIOLOGICAL PROCESSES**	**MOLECULAR FUNCTION**	**CELLULAR COMPONENT**	**DESCRIPTION**
>EcFEMNSOD1	GO: 0019430	GO: 0004784	GO: 0005739	Superoxide dismutase
>EcFEMNSOD2	GO: 0019430	GO: 0004784	GO: 0005739	Superoxide dismutase
>EcCUZNSOD1	GO: 0006801	GO: 0004784	GO: 0005615	Superoxide dismutase (Cu-Zn)
>EcCUZNSOD2	GO: 0006801	GO: 0005507		Superoxide dismutase (Cu-Zn)
>EcCUZNSOD3	GO: 0006801	GO: 0004784	GO: 0005507	Superoxide dismutase (Cu-Zn)
>EcCUZNSOD4	GO: 0006801	GO: 0005507	GO: 0005507	Superoxide dismutase (Cu-Zn)
>EcCUZNSOD5	GO: 0006801	GO: 0004784	GO: 0005615	Superoxide dismutase (Cu-Zn)
>EcCUZNSOD6	GO: 0006801	GO: 0004784	GO: 0005615	Superoxide dismutase (Cu-Zn)
>EcCUZNSOD7	GO: 0006801	GO: 0005507		Superoxide dismutase (Cu-Zn)
>EcCUZNSOD8	GO: 0006801	GO: 0004784	GO: 0005615	Superoxide dismutase (Cu-Zn)

Table 6 – CpG islands predicted in promoter and gene body sequences.

Gene ID	No of islands	Region
EcCUZNSOD1	0	Promoter
	2	Gene body
EcCUZNSOD2	3	Promoter
	1	Gene body
EcCUZNSOD3	2	Promoter
	1	Gene body
EcCUZNSOD4	0	Promoter
	1	Gene body
EcCUZNSOD5	3	Promoter
	0	Gene body
EcCUZNSOD6	0	Promoter
	2	Gene body
EcCUZNSOD7	4	Promoter
	1	Gene body
EcCUZNSOD8	3	Promoter
	1	Gene body
EcFeMnSOD1	3	Promoter
	1	Gene body
EcFeMnSOD2	2	Promoter
	1	Gene body

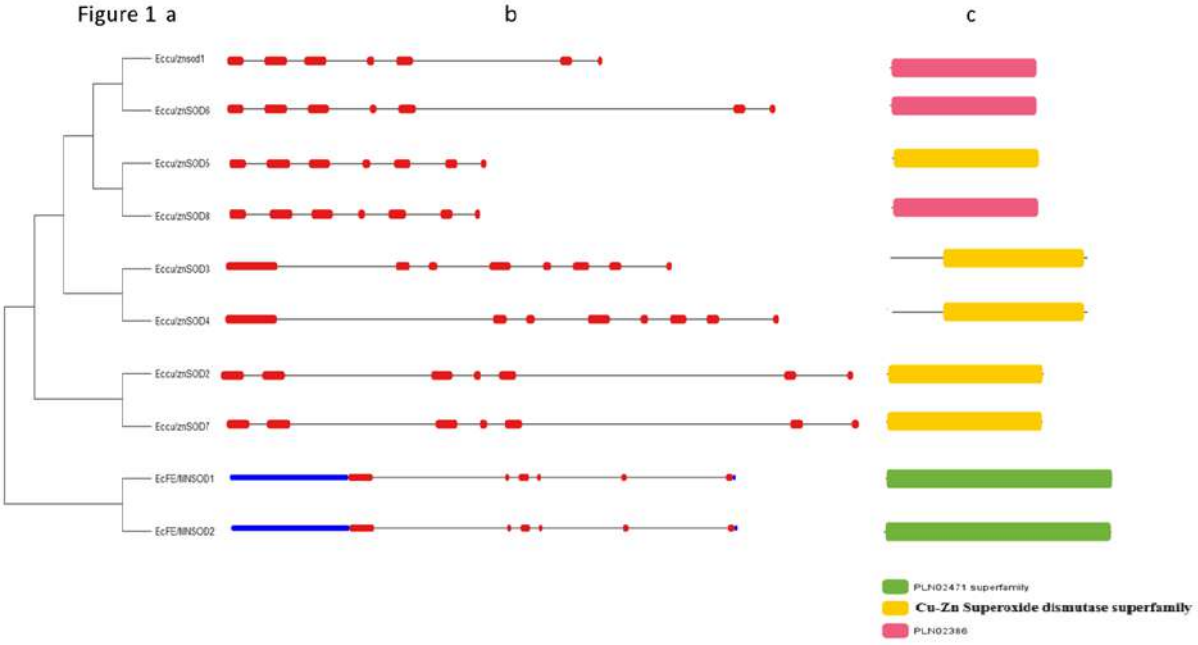


Fig.1- a) Phylogenetic analysis of EcSOD homologs in tree form. b) Distribution pattern of exon/intron of candidate

genes where red colour = exons, blue = UTRs and straight lines = introns. c) Domain analysis done by NCBI-CDD search.

Figure 2

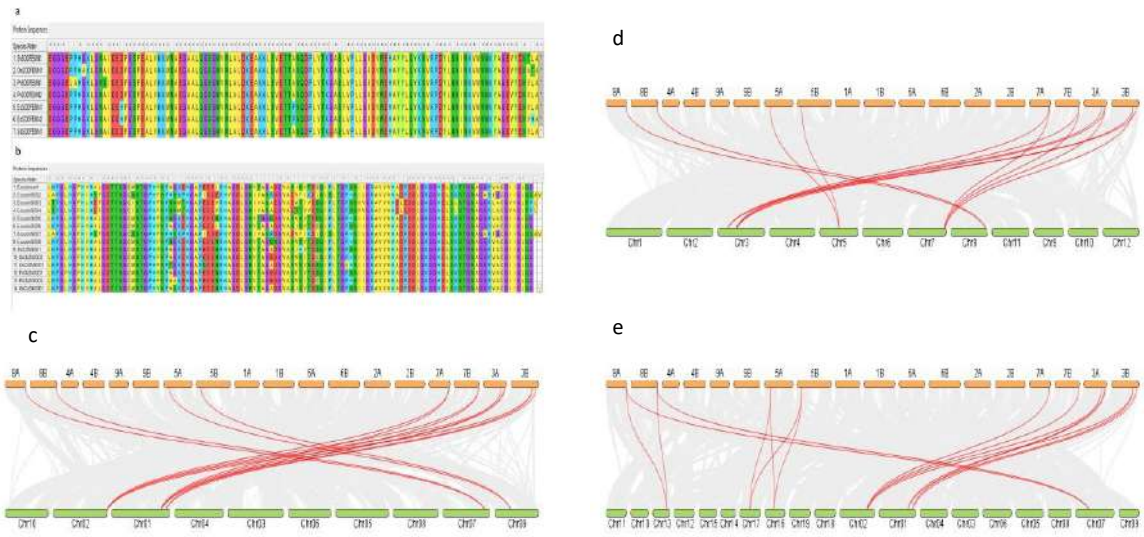


Fig.2- a) Multiple sequence analysis of Fe/Mn SODs among related species. b) Multiple sequence analysis of Cu/Zn SODs among related species. c) Synteny analysis of E. corcana vs S. bicolor. d) Synteny analysis of E. corcana vs O. sativa. e) Synteny analysis of E. corcana vs M. sinesis where pink lines indicate collinearity of candidate EcSOD genes and grey colour indicates collinearity among genomes.

Figure 3

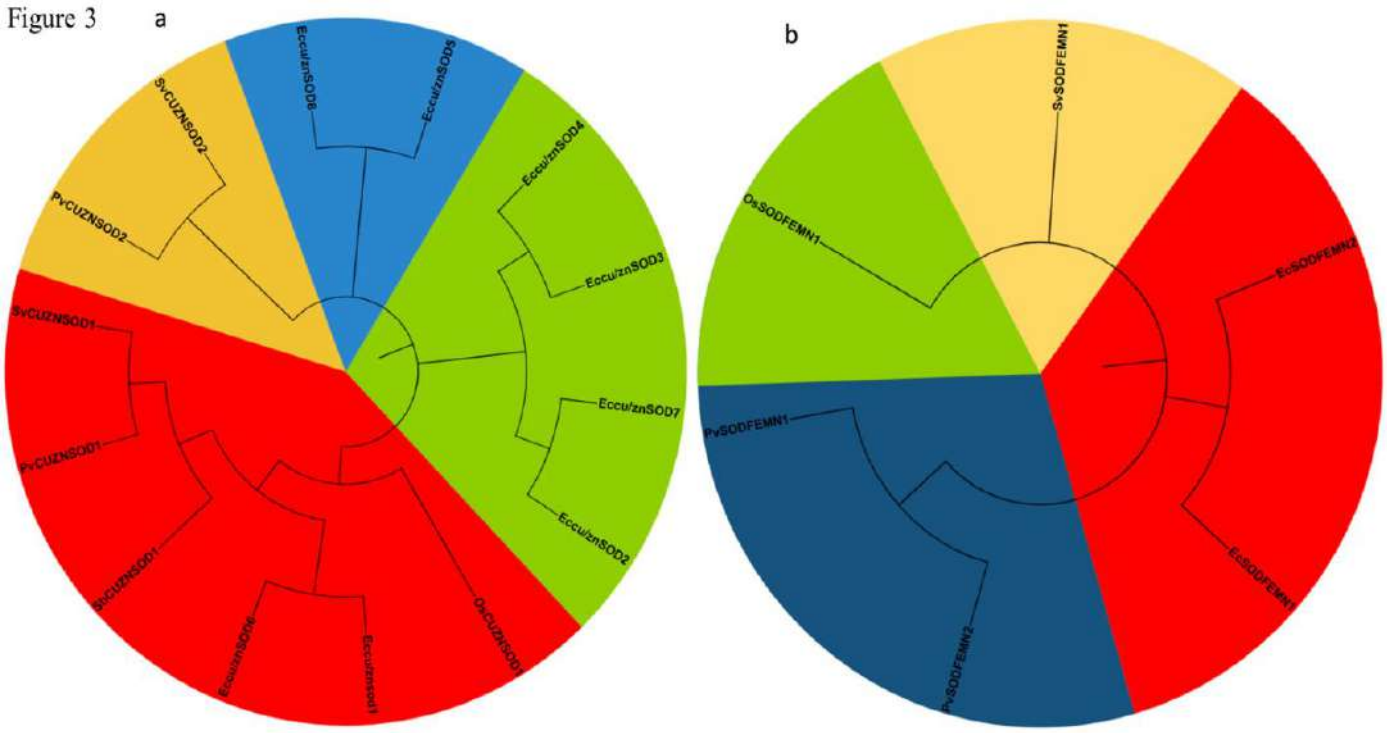


Fig.3 – a) Evolutionary relationship of EcCUZNSOD homolog genes with other species like Sorghum bicolor, Oryza sativa, Panicum virgatum, and Seteria viridis. b) Evolutionary relationship of EcFEMNSOD homolog genes with other species like Sorghum bicolor, Oryza sativa, Panicum virgatum, and Seteria viridis.

Figure 4

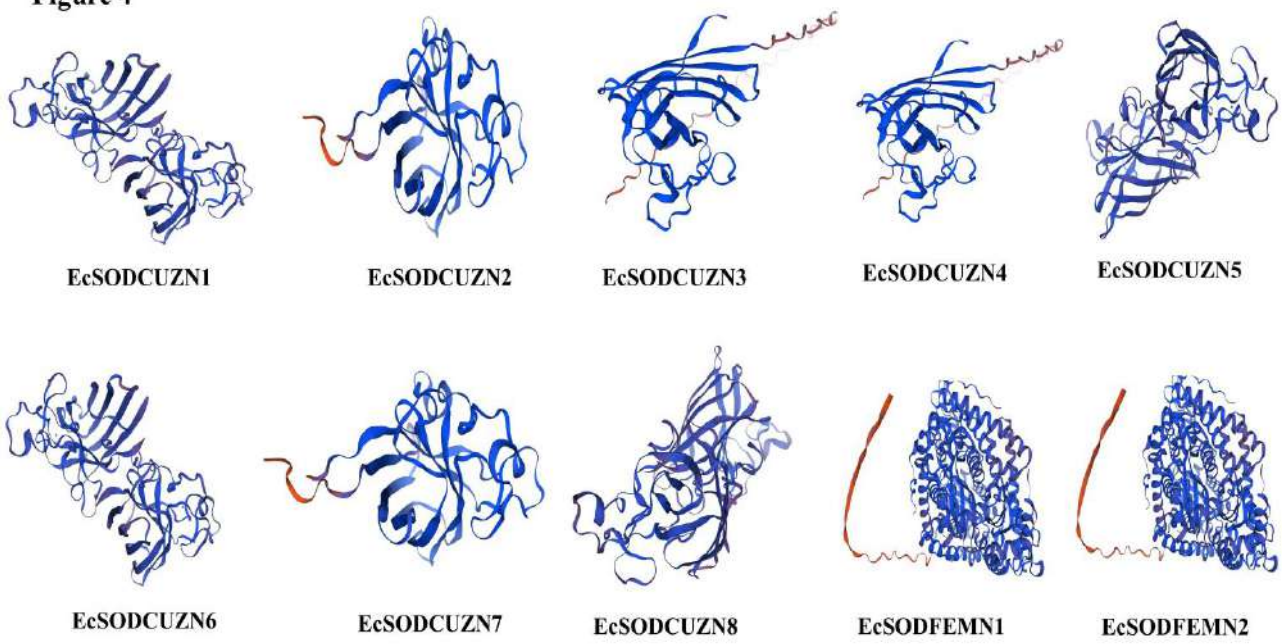


Fig.4 – 3D structure analysis of EcSODs proteins.

Figure 5

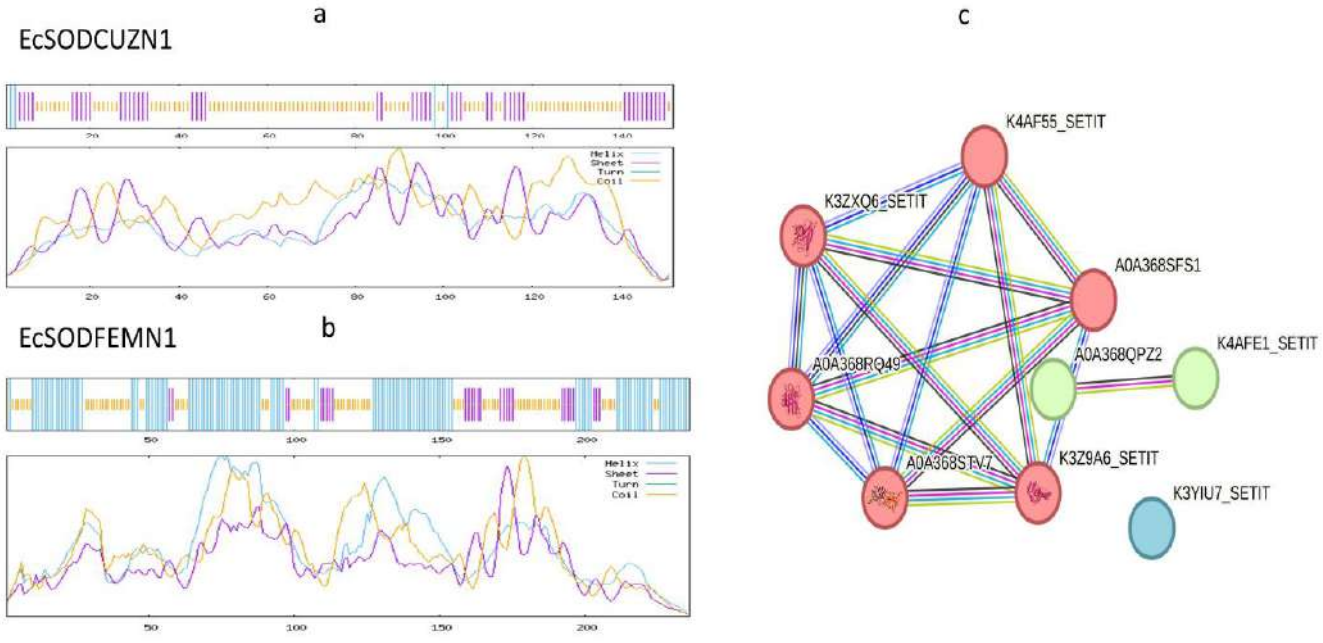


Fig.5 – a and b) secondary structure of EcSODCUZN1 and EcFEMNSOD1 proteins respectively. c) Protein -protein interactions among EcSOD candidate proteins selected.

Figure 6

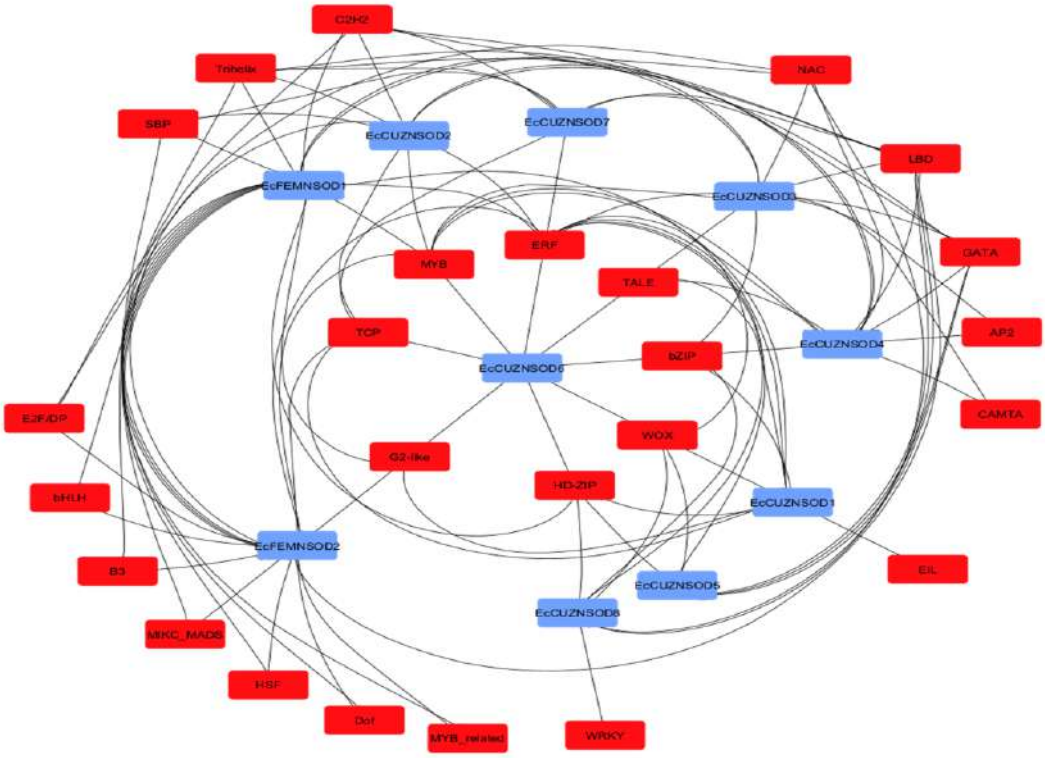


Fig.6 – Network created by Cytoscape tool shows predicted Transcriptional factors.

Figure 7

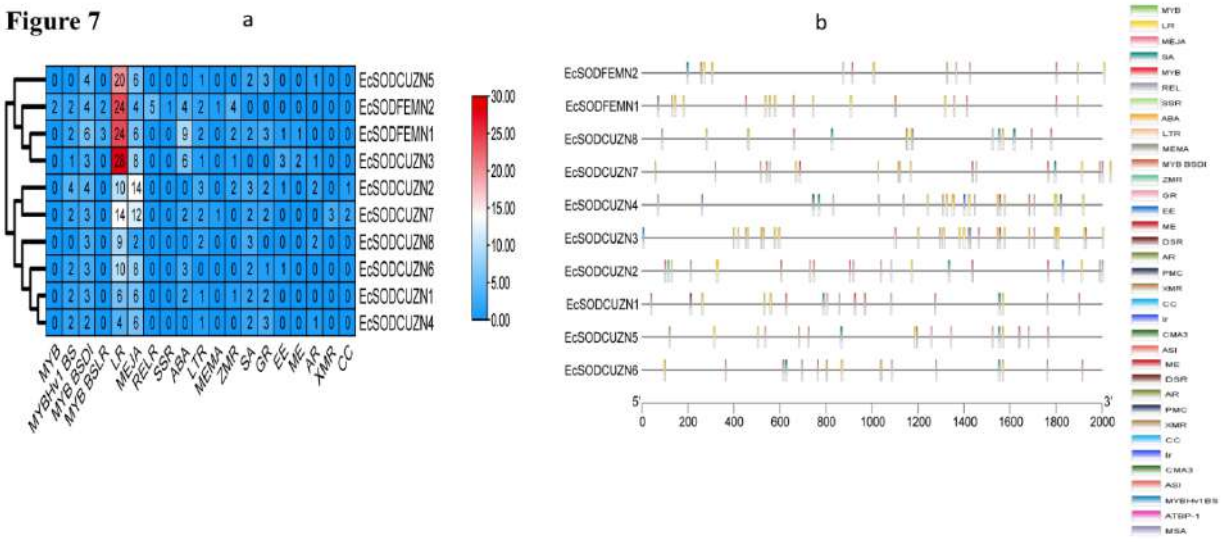


Fig.7- a) Heatmap representing distribution of cis regulatory elements among EcSOD homolog genes b) The distribution of promotor elements predicted.

Figure 8

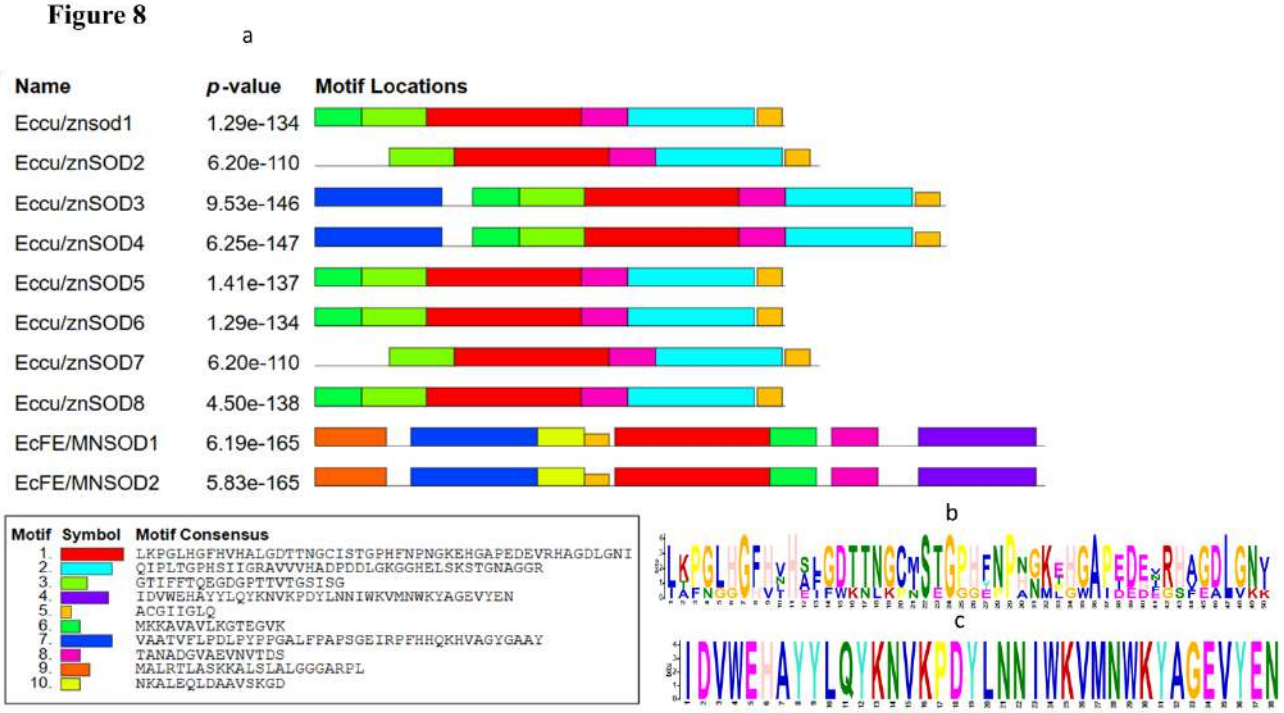


Fig.8- a) Motif structure analysis by MEME tool. Red colour motif is Superoxide dismutase activity. Blue colour motif is SOD with Cu/Zn ligands. Violet colour motif is SOD with Fe/Mn metal ligands. b) Motif pattern of SOD (red colour motif). c) Motif sequence pattern of Fe/Mn violet colour motif.

Figure 9

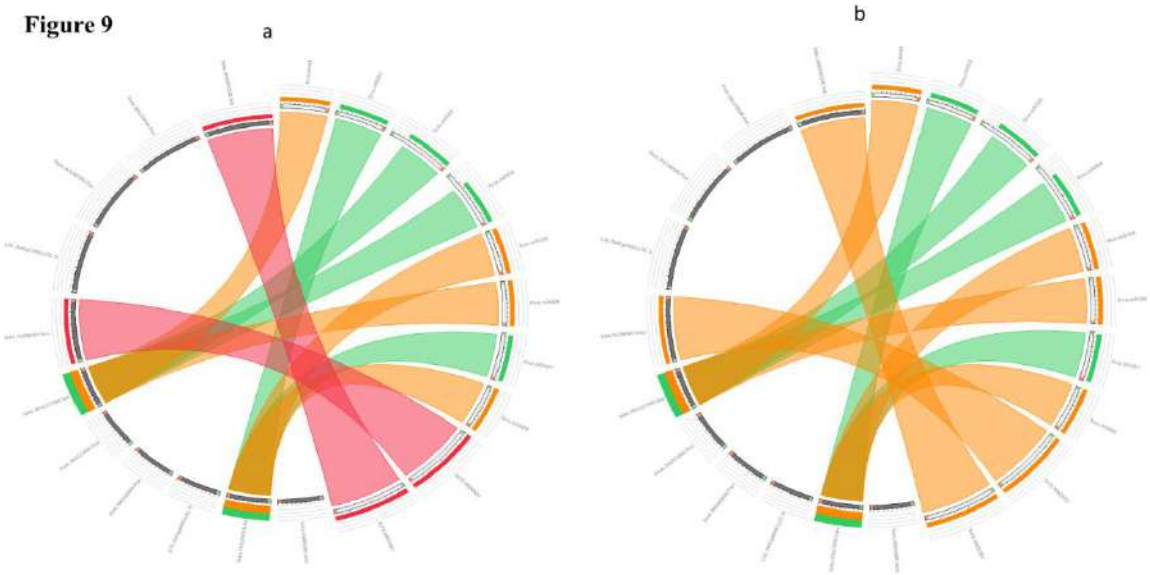


Fig.9- Multiple circos similarity analysis by Circoletto tool. a) analysis done through bitscore values where red colour > 0.75, orange = 0.75 and blue <0.75 value. b) Analysis done through %identity of protein sequences where orange colour means higher identity i.e. 99.999% and blue colour less than 99.9% similarity among species.

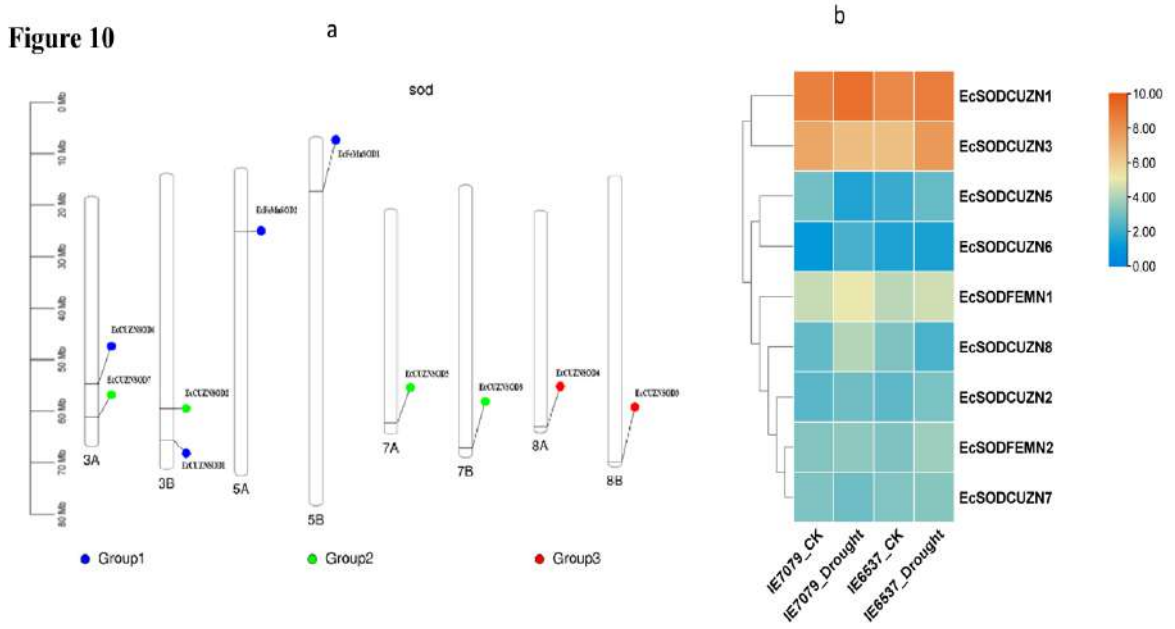


Fig.10- a) Phenogram of EcSODs genes on their respective chromosomes. b) Heatmap showing the expression levels of EcSODs genes based on TKM values.

V. CONCLUSION

Using a wide variety of bioinformatic methods, we have examined the finger millet genome in this work in order to discover and define the *SOD* gene family. *SOD* are the main antioxidants and are among those initial to take part in the ROS species scavenging process. It plays a key role in understanding how plants counter to stress. The findings of this work provide support for the functional characterization of *EcSOD* proteins and deepen our knowledge of the evolutionary links within the *SOD* family. To sum up, our research has yielded extensive details regarding the ten *SOD* genes found in finger millet, such as gene structures, phylogenetic relationships, chromosome locations as well as gene ontology. These important findings suggest that *EcSOD* genes play a significant role in controlling the development of plant tissue and are probably involved in the response to both abiotic and biotic stress. This methodical identification of the whole genome offers a foundation for further research on the role of *EcSOD* proteins in biological processes. It also offers a possible means of improving finger millet breeding under various biotic and abiotic stress, as well as for investigations including gene editing and manipulation.

FUNDING

No additional funding is provided for this work.

ACKNOWLEDGEMENTS

Viswanadha Naik Jarapala is thankful to University Grants Commission (UGC) for his research fellowship (211610001184) and thankful to Centre for Plant Molecular Biology (CPMB) & Department of Genetics and Biotechnology, Osmania University-Hyderabad.

AUTHOR CONTRIBUTION STATEMENT

Viswanadha Naik. J and Srinivas Naik. K came up with the experiments and the article's framework. Viswanadha Naik. J and Anjana priyadarshani. K wrote the initial draft. All others have contributed lateral text to the manuscript and improved it. Prashanth B, Viswanadha Naik. J and Vikas Reddy has provided the figures. Srinivas Naik. K, Viswanadha Naik. J, Prashanth. B, Anjana priyadarshani. K, Vikas Reddy. O and Vijay Kumar. G edited the manuscript. All the Authors have given their approval.

STATEMENTS AND DECLARATIONS

Declaration of competing interest

The authors declare that they have no known conflicts of interest.

Data availability statement

The complete data is accessible with the corresponding author K. Srinivas Naik.

REFERENCES

[1] Kumari P, Thakur A, Sankhyan KN, Singh U (2023) Millet production and consumption in india and their nutritional aspects. Just Agric 3:236–243

[2] Meena RP, Joshi D, Bisht JK, Kant L (2021) Global Scenario of Millets Cultivation. In: Millets and Millet Technology. Springer Singapore, pp 33–50

[3] Rathore T, Singh R, Kamble DB et al (2019) Review on finger millet: processing and value addition. Pharma Innov J 8:283–291

[4] Tadele, Zerihun. "Orphan crops: their importance and the urgency of improvement." Planta 250 (2019): 677-694.

[5] Kumar A, Metwal M, Kaur S, Gupta AK, Puranik S, Singh S et al. Nutraceutical value of finger millet [*Eleusine coracana* (L.) Gaertn.], and their improvement using omics approaches. Front Plant Sci. 2016;7:934.

[6] Maharajan T, Ceasar SA, Ajeesh Krishna TP. Finger Millet (*Eleusine coracana* (L.) Gaertn): nutritional importance and nutrient transporters. Crit Rev Plant Sci. 2022;41:1–31.

[7] Wambi W, Otienno G, Tumwesigye W, Mulumba J. Genetic and genomic resources for finger millet improvement: opportunities for advancing climate-smart agriculture. J Crop Improv. 2021;35:204–33.

[8] Devi PB, Vijayabharathi R, Sathyabama S et al (2014) Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. J Food Sci Technol 51:1021–1040

[9] Jia, Xin, et al. "Comparative Transcriptome Analysis Reveals Molecular Mechanism and Regulatory Network of Rosa Chinensis 'Old Blush'Leaves in Response to Drought Stress." (2020).

[10] Mareri, Lavinia, Luigi Parrotta, and Giampiero Cai. "Environmental stress and plants." International Journal of Molecular Sciences 23.10 (2022): 5416.

[11] Ramesh, Gutha Venkata, et al. "Breeding finger millet for biotic stress resistance." Genetic improvement of Small Millets. Singapore: Springer Nature Singapore, 2024. 279-295.

[12] Yoshida, Kentaro, et al. "Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable elements." *BMC genomics* 17 (2016): 1-18.

[13] Akech, V., Ojulong, H., Okori, P., & Biruma, M. (2016). Resistance to blast in interspecific finger millet progenies. fifth RUFORUM biennial regional conference 17-21 October 2016, Cape Town, *South Africa*. RUFORUM working document series, 14, 725–730.

[14] Nagaraja, A., Jagadish, P. S., Ashok, E. G., & Gowda, K. T. (2007). Avoidance of finger millet blast by ideal sowing time and assessment of varietal performance under rain fed production situations in Karnataka. *Journal of Mycopathological. Research*, 45(2), 237–240.

[15] Mbinda, W., & Masaki, H. (2021). Breeding strategies and challenges in the improvement of blast disease resistance in finger millet. a current review. *Frontiers in Plant Science*, 11, 602882. https://doi.org/10.3389/fpls.2020.602882

[16] Gechev TS, Van Breusegem F, Stone JM et al (2006) Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. BioEssays 28:1091–1101

[17] Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399

[18] Sugimoto M, Oono Y, Gusev O et al (2014) Genome-wide expression analysis of reactive oxygen species gene network in Mizuna plants grown in long-term spaceflight. BMC Plant Biol 14:4

[19] Tepperman JM, Dunsmuir P (1990) Transformed plants with elevated levels of chloroplastic SOD are not more resistant to superoxide toxicity. Plant Mol Biol 14:501–511

[20] Gopavajhula VR, Chaitanya KV, Khan PAA, Shaik JP, Reddy PN and Alanazi M (2013) Modeling and analysis of soybean (Glycine max. L) Cu/Zn, Mn and Fe superoxide dismutases. Genet Mol Biol 36:225-236

[21] Fink, Ryan C., and John G. Scandalios. "Molecular evolution and structure–function relationships of the superoxide dismutase gene families in angiosperms and their relationship to other eukaryotic and prokaryotic superoxide dismutases." Archives of Biochemistry and Biophysics 399.1 (2002): 19-36.

[22] Dupont CL, Neupane K, Shearer J, Palenik B. Diversity, function and evolution of genes coding for putative Ni-containing superoxide dismutases. *Environmental Microbiology.* 2008;10(7):1831–1843. doi: 10.1111/j.1462-2920.2008.01604.x. [PubMed] [CrossRef] [Google Scholar] [Ref list]

[23] Verma, D., Lakhanpal, N. & Singh, K. Genome-wide identification and characterization of abiotic-stress responsive *SOD* (*superoxide dismutase*) gene family in *Brassica juncea* and *B. rapa*. *BMC Genomics* 20, 227 (2019). https://doi.org/10.1186/s12864-019-5593-5

[24] Tutar O, Marín-Guirao L, Ruiz J, Procaccini G. Antioxidant response to heat stress in seagrasses. A gene expression study. *Marine Environmental Research.* 2017;132:94–102. doi: 10.1016/j.marenvres.2017.10.011.

[25] Winters G, Nelle P, Fricke B, Rauch G, Reusch TBH. Effects of a simulated heat wave on photophysiology and gene expression of high- and low-latitude populations of *Zostera marina*. *Marine Ecology Progress Series.* 2011;435:83–95. doi: 10.3354/meps09213.

[26] R. Madanala, V. Gupta, F. Deeba, S.K. Upadhyay, V. Pandey , P.K. Singh, R. TuliBiotechnology letters, 33 (2011), pp. 2057-2063

[27] S.S. Gill, N.A. Anjum, R. Gill, S. Yadav, M. Hasanuzzaman, M. Fujita, N. Tuteja Superoxide dismutase—mentor of abiotic stress tolerance in crop plants Environ. Sci. Pollut. Res., 22 (2015), pp. 10375-

[28] Q. Guan, X. Liao, M. He, X. Li, Z. Wang, H. Ma, S. Liu Tolerance analysis of chloroplast OsCu/Zn-SOD overexpressing rice under NaCl and $NaHCO_3$ stress PLoS ONE, 12 (10) (2017), Article e0186052

[29] Wang et al. (2018b) Wang L, Wang L, Zhang Z, Ma M, Wang R, Qian M, Zhang S. Genome-wide identification and comparative analysis of the superoxide dismutase gene family in pear and their functions during fruit ripening. *Postharvest Biology and*

Technology. 2018b;143:68–77.
doi: 10.1016/j.postharvbio.2018.04.012

[30] David, R. Host Antony, et al. "Mining QTL and genes for root traits and biochemical parameters under vegetative drought in South Indian genotypes of finger millet (Eleusine coracana (L.) Gaertn) by association mapping and in silico comparative genomics." *Biocatalysis and Agricultural Biotechnology* 32 (2021): 101935.

[31] Devos, Katrien M., et al. "Genome analyses reveal population structure and a purple stigma color gene candidate in finger millet." *Nature communications* 14.1 (2023): 3694.

[32] Hatakeyama, Masaomi, et al. "Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop." *DNA Research* 25.1 (2018): 39-47.

[33] Goodstein DM, Shu S, Howson R et al (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40:D1178–D1186. https://doi.org/10.1093/nar/gkr944

[34] Hu B, Jin J, Guo A-Y et al (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296–1297. https://doi.org/10.1093/bioinformatics/btu817

[35] Bailey TL, Boden M, Buske FA et al (2009) MEME suite: tools for motif discovery and searching. Nucleic Acids Res 37:202–208. https://doi.org/10.1093/nar/gkp335

[36] Horton P, Park K-JK-J, Obayashi T et al (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Res 35:W585–W587. https://doi.org/10.1093/nar/gkm259.

[37] Geourjon, Christophe, and Gilbert Deleage. "SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments." Bioinformatics 11.6 (1995): 681-684.

[38] Liu, Z., Fan, M., Li, C., & Xu, J. H. (2019). Dynamic gene amplification and function diversification of grass-specific O-methyltransferase gene family. Genomics, 111(4), 687-695.

[39] Xiao, Peixuan, et al. "Evolution analyses of CAMTA transcription factor in plants and its enhancing effect on cold-Tolerance." Frontiers in Plant Science 12 (2021): 758187.

[40] Han, Guoliang, et al. "C2H2 zinc finger proteins: master regulators of abiotic stress responses in plants." Frontiers in plant science 11 (2020): 115

[41] Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399

[42] Foyer, Christine H., Alexander V. Ruban, and Graham Noctor. "Viewing oxidative stress through the lens of oxidative signalling rather than damage." *Biochemical Journal* 474.6 (2017): 877-883.

[43] Perry, J. J. P., et al. "The structural biochemistry of the superoxide dismutases." *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1804.2 (2010): 245-262.

[44] Zhang H, Xu N, Li X, Gu S. 2016. Overexpression of 2-Cys Prx increased salt tolerance of photosystem II(PSII) in tobacco. PeerJ Preprints 4:e2500v1. https://doi.org/10.7287/peerj.preprints.2500v1.

[45] Wang, Wei, et al. "Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in Gossypium raimondii and G. arboreum." Data in brief 9 (2016): 345-348.

[46] Wang LB, Wang L, Zhang Z et al (2018) Genome-wide identification and comparative analysis of the superoxide dismutase gene family in pear and their functions during fruit ripening. Postharvest Biol Tec 143:68–77

[47] Song J, Zeng L, Chen R et al (2018) In silico identification and expression analysis of superoxide dismutase (*SOD*) gene family in *Medicago truncatula*. 3 Biotech 8(8):348

[48] Xu G, Guo C, Shan H et al (2012) Divergence of duplicate genes in exon-intron structure. Proc Natl Acad Sci USA 109:1187–1192

[49] Cheng H, Song S, Xiao L et al (2009) Gibberellin acts through jasmonate to control the expression of MYB. PLoS Genet 5 21:e1000440 M M t Arabidopsis

[50] Salih H, Gong W, He S et al (2016) Genome-wide characterization and expression analysis of MYB transcription factors in Gossypium hirsutum. BMC Genet 17:129

[51] Biswas, Dew, Hena Gain, and Arunava Mandal. "MYB transcription factors in biotic stress tolerance." Plant Stress (2023): 100252.

[52] Vom Endt, Debora, Jan W. Kijne, and Johan Memelink. "Transcription factors controlling plant secondary metabolism: what regulates the regulators?." Phytochemistry 61.2 (2002): 107-114.

[53] Seifikalhor, Maryam, et al. "Calcium signaling and salt tolerance are diversely entwined in plants." Plant Signaling&Behavior 14.11 (2019): 1665455.

[54] Li, Long-Cheng, and Rajvir Dahiya. "MethPrimer: designing primers for methylation PCRs." *Bioinformatics (Oxford, England)* vol. 18,11 (2002): 1427-31. doi:10.1093/bioinformatics/18.11.1427